4Д

УДК 004.89, 004.93

А.В. Дьяченко
Институт проблем искусственного интеллекта МОН Украины и НАН Украины, г. Донецк
nastya-djachenko@yandex.ru

Задача 3D распознавания лиц: современные методы решения

В статье рассмотрено современное состояние проблемы 3D распознавания лиц, приведены принципы работы существующих систем трехмерного распознавания. Представлены способы решения ключевой задачи в области трехмерного распознавания – получения трехмерной информации о лице. Рассмотрены распространенные методы 3D распознавания лиц.

Введение

Главной задачей в области распознавания зрительных образов по-прежнему остается распознавание лиц. Множество исследований и разработок посвящено этой проблеме, однако эффективность существующих систем распознавания лиц еще далека от возможностей человека.

В настоящее время актуальным и интенсивно разрабатываемым направлением исследований в области распознавания зрительных образов, в особенности лиц, является направление, связанное с получением трехмерной информации об объекте.

Алгоритмы трехмерного распознавания используют информацию о глубине и кривизне поверхности, в отличие от систем двухмерного распознавания, традиционно использующих признаки, основанные на яркости пикселей изображения. Следовательно, трехмерные дескрипторы имеют большую точность в описании особенностей поверхности; лучше подходят для описания свойств лица в областях щек, лба и подбородка; инвариантны к ракурсу и освещению.

На данный момент существует задача создания системы распознавания лиц на основе извлечения 3D информации из видеопотока, состоящая в следующем: человек последовательно поворачивает голову по трем степеням свободы (наклоны вперед-назад и влево-вправо, повороты влево-вправо) перед камерой, по полученной последовательности кадров строится трехмерная модель головы и производится распознавание лица (сопоставление полученной модели с имеющимися моделями в базе данных).

Целью данной статьи является рассмотрение современных методов 3D распознавания лиц для решения поставленной задачи разработки эффективной системы распознавания лиц.

Современное состояние проблемы 3D распознавания лиц

Существующие на данный момент системы 3D распознавания лиц используют специальное оборудование для реконструкции трехмерной модели лица (сенсорные системы). Сенсорные технологии 3D распознавания делятся на три категории:

1. Стерео. Используются две камеры с известным взаиморасположением для получения стереопары изображений объекта; на полученных изображениях находятся соответствующие точки и вычисляется положение сопоставленных точек в трехмерном пространстве.

2. Структурированный свет. Этот подход использует камеру и световой проектор: структурный свет проецирует на лицо специальную текстуру, а камера регистрирует

искажения этой текстуры на объемном объекте. С помощью методов восстановления формы по текстуре вычисляется расположение точек в трехмерном пространстве.

3. Лазерное сканирование. Лазерные сканеры применяют свет как источник для обнаружения расстояния до объекта сканирования. Они измеряют время отражения лазера от объекта и получают информацию о глубине расположения точек на его поверхности.

Несмотря на то, что подобные технологии дают очень высокий результат (ошибка распознавания в пределах одного процента), даже при идеальном освещении сенсорные системы подвержены недостаткам: частым случаем является появление на модели артефактов в виде «дыр» и «выступов» из-за отсутствующих данных и ошибок восстановления. Другим недостатком сенсорных систем является маленькая глубина резкости для получения необходимой информации, для стереосистем – около 0,3 м, для систем со структурированной подсветкой – порядка одного метра. Еще одним недостатком является сложное и зачастую дорогостоящее оборудование.

На сегодняшний день можно отметить следующие компании, занимающиеся развитием технологии сенсорного 3D распознавания: Geometrix (США), Genex Technologies (США), Bioscrypt (Канада), L-1 Identity Solutions (Англия). В России в данном направлении работает компания Artec Group.

Системы 3D распознавания лиц, не использующие дополнительное оборудование, существуют только в качестве опытных разработок и коммерческого применения пока не имеют.

Способы получения трехмерной информации о лице

Для того чтобы получить трехмерную информацию об объекте, используются алгоритмы, объединенные в англоязычной литературе под названием «shape from X» (получение формы с помощью X), где под X понимаются самые различные методы. Рассмотрим те из них, которые наиболее перспективны с точки зрения решения поставленной во введении задачи.

Восстановление формы по теням (shape from shading, SFS). Задача восстановления формы объекта по изменению яркости пикселей изображения основана на способности человека определить форму объекта, используя визуальную информацию о характере отражения света на его поверхности. Данная задача является задачей, обратной задаче визуализации (рендеринга), когда яркость точки моделируемой сцены зависит от ряда факторов и вычисляется согласно заданной математической модели освещения [1].

Среди факторов, влияющих на яркость точки поверхности объекта, можно выделить следующие:

1. Свойства и расположение источников света.
2. Характеристики поверхности, определяющие ее отражающие свойства.
3. Ориентация участка поверхности, соответствующего данной точке в пространстве.
4. Точка зрения наблюдателя.

В качестве математической модели взаимодействия света и поверхности обычно используется модель рассеяния Ламберта, описывающая функцию зависимости яркости точки изображения от интенсивности единственного источника света, альбедо (коэффициента отражения) поверхности и скалярного произведения единичной нормали к поверхности и вектора направления на источник света.

Поскольку данная математическая модель содержит большое количество неизвестных параметров, чтобы свести задачу SFS к решаемому виду, применяются различные упрощения, в основном относительно направления освещения.

Соответственно недостатками алгоритмов «shape from shading» являются: необходимость априорных знаний о законе рассеяния, слишком общие допущения об отражательных свойствах поверхности, приводящие к неправильному восстановлению формы поверхности. Наиболее полно проблема восстановления формы по закраске с точки зрения распознавания лиц освещена в [2].

Восстановление формы по стереопаре (shape from stereo). Построение трёхмерных моделей по стереопаре традиционно рассматривается как две последовательные задачи: стереосопоставление и построение трёхмерных моделей по множеству точек. Задачей алгоритма является получение данных о расстоянии до объектов сцены, на основании которых строится карта смещений (disparity map).

Большинство существующих алгоритмов стереосопоставления может быть разбито на две категории решений: локальную и глобальную. Локальные методы основаны на поиске характеристических точек (feature points) и их сопоставлении между двумя кадрами. Глобальные методы ищут соответствие между изображениями по каждой паре пикселей, а поскольку бывают области, где какая-либо текстура отсутствует, то налагаются ограничения гладкости. Хорошие алгоритмы ищут карту смещений как кусочно-гладкую функцию, с ограниченным количеством линий разрыва и учитывают, что некоторые точки видны только на одном изображении. Вообще говоря, локальные алгоритмы являются с точки зрения вычислений менее затратными, а глобальные алгоритмы формируют более точные карты смещений.

Восстановление трехмерной сцены по стереопаре способно дать качественный результат и восстановить трехмерный образ практически для всех точек исходного изображения, однако требует высокой точности калибровки камер стереопары.

При недостатке пространственной структуры объекта (отсутствии ярко выраженных характерных точек и текстуры) алгоритмы стереовосстановления находят только грубые детали формы объекта.

Восстановление формы по движению (shape from motion, SFM). Данный метод реконструкции трехмерных сцен использует относительное движение между камерой и сценой в последовательности изображений [3]. Как и в стереовосстановлении, задачу SFM можно разделить на две подзадачи: нахождение взаимно однозначного соответствия характеристических точек на последовательных кадрах и реконструкция сцены. Но есть и некоторые важные различия. Разница между последовательными кадрами намного меньше, чем между изображениями в типичной стереопаре, поскольку видео снимается с частотой несколько десятков кадров в секунду. Также, в отличие от стерео, в движении относительное смещение между камерой и сценой не обязательно вызвано одинаковым трехмерным преобразованием.

Что касается поиска соответствий, то алгоритм SFM предоставляет множество тесно связанных изображений (кадров видео) для анализа, и это является преимуществом данного подхода. Во-первых, здесь могут быть использованы методы отслеживания, которые используют историю движений для предсказания различий в следующем кадре. Во-вторых, проблема соответствия также может быть рассмотрена как задача оценки видимого движения на изображении (оптический поток).

Для определения соответствий используется, как правило, два вида методов. Дифференциальные методы применяют оценки производных по времени и поэтому требуют плотную выборку последовательных изображений. Этот метод работает с каждым пикселем изображения и приводит к плотным измерениям. Другие методы используют фильтр Калмана для сопоставления и отслеживания точечных характеристик. Эти методы работают с небольшим количеством точек изображения и приводят к разреженным измерениям.

В отличие от задачи нахождения соответствий, задача реконструкции в данном подходе более сложна по сравнению со стереовосстановлением. Восстановление движения и структуры кадр за кадром оказывается более чувствительным к шуму. Причина в том, что базовая линия между последовательными кадрами очень мала.

В задаче восстановления формы по движению применяются алгоритмы факторизации матриц, с помощью которых можно восстановить положение и ориентацию камер, внутренние параметры камер (фокусное расстояние), т.е. параметры, которые очень часто неизвестны. Кроме того, большое количество кадров дает возможность проверки правильности сопоставления. Дополнительно в ряде случаев можно еще и получать оценки точности восстановления, соответствующие именно данной сцене.

Методы SFM, основанные на факторизации матриц, не работают непосредственно с изображениями, а требуют на вход координаты характеристических точек изображений в пикселях и наличие у каждой характеристической точки маркера (номера), причем на всей последовательности изображений одной и той же точке реальной сцены должен соответствовать одинаковый маркер. Таким образом, восстанавливаются трехмерные координаты только некоторых точек сцены, поэтому возникает задача интерполяции поверхностей сцены между восстановленными точками модели.

Алгоритмы трехмерного распознавания лиц

Среди различных подходов 3D распознавания можно выделить три основных: анализ формы 3D поверхности лица, статистические подходы и использование параметрической модели лица [4].

Анализ формы 3D поверхности. Методы, основанные на анализе формы трехмерного изображения лица, используют локальные или глобальные характеристики поверхности, которая описывает лицо, например, кривизну, профили линий, метрики расстояний между двумя поверхностями.

Кривизна поверхности используется для сегментации поверхности лица по признакам, которые могут быть использованы для сравнения поверхностей. Другой подход основан на 3D дескрипторах поверхности лица в терминах средней и гауссовой кривизны либо в терминах расстояний и отношении углов между характерными точками поверхностей. Еще одним локально-ориентированным методом является подход, использующий точки-сигнатуры. Идея метода заключается в формировании представления-описания выбранной точки по соседним точкам вокруг заданной точки поверхности. Эти сигнатуры точек используются для сравнения поверхностей.

Глобальные методы задействуют всю информацию о трехмерном изображении лица как входную для системы распознавания. Например, модель лица выравнивается на основе его зеркальной симметрии, после чего выделяются и сравниваются профили лица вдоль плоскости выравнивания. Также используется метод сравнения моделей лиц на основе максимального и минимального значений и направления кривизны профилей.

Еще один подход основан на методе сравнения расстояний между поверхностями для распознавания. Одни методы основаны на вычислении метрик наименьших расстояний между поверхностями моделей, другие – на измерениях расстояния не только между поверхностями, но и текстурой на поверхности. Однако существенным ограничением данных методов является то, что лицо не должно деформироваться и его поверхность является жесткой.

Третий подход берет за основу извлечение и анализ трехмерных профилей и контуров, выделенных на лице.

Существуют еще и гибридные методы, основанные на объединении локальной информации о поверхности в виде локальных моментов с глобальной трехмерной сеткой, описывающей поверхность всего лица. В одном из таких методов значение функции Z(x,y), описывающей карту глубины лица в выровненной системе координат, разлагается на Фурье-компоненты. Разложение функции на моменты (базисные функции) позволяет сгладить мелкий высокочастотный «шум на лице» и случайные выбросы. Кроме Фурье-разложения используются и другие базисные функции: степенные ряды, полиномы Лежандра и моменты Цернике.

Статистические методы, в частности метод главных компонент (Principal Component Analysis, PCA), ранее широко использовались в 2D распознавании. Метод PCA реализован также для 3D распознавания и был расширен одновременно на комбинацию карт глубины и цвета. Альтернативным для PCA является метод линейного дискриминантного анализа, в котором, в отличие от PCA, один объект (заданный человек) задается не одним лицом, а набором моделей (3D лиц).

Еще одним статистическим методом, также позаимствованным из 2D распознавания, является метод скрытых марковских моделей (СММ). Теория Марковских случайных полей позволяет строить оценки различных пространственно-переменных величин по изображениям, при этом накладывая на эти величины определенные априорные ограничения. В качестве таких пространственно-переменных величин могут выступать, например, значения смещения в задаче стереореконструкции. В литературе по 3D распознаванию этот метод известен как псевдо 3D скрытые марковские модели и используется, в частности, для распознавания выражений лиц.

Использование параметрической модели лица. Ключевая идея распознавания по моделям основана на так называемых параметрических 3D моделях, когда форма лица контролируется набором параметров (коэффициентов) модели. Эти коэффициенты описывают 3D форму лица и также могут задавать цвет (текстуру) на ее поверхности. Данный метод использует на входе одно или несколько изображений лица, в основном фотографий, полученных в анфас и профиль [5].

Алгоритм решения задачи строится по итеративному принципу. В качестве исходной итерации выбирается некоторый усредненный трехмерный макет головы человека, и производится пошаговое его улучшение. При этом используется набор антропометрических точек лица, извлеченных из фотографии, который деформируется до заданной трехмерной поверхности. Параметры деформации вычисляются в процессе 3D реконструкции с помощью эластичной модели. Эти параметры потом используются для распознавания в качестве вектора признаков данного лица.

Заключение

Рассмотренные выше способы восстановления трехмерной модели лица имеют ограничения и недостатки, которые не позволяют с их помощью решить задачу распознавания лиц по видеопоследовательности эффективно. Методы восстановления формы по движению, основанные на факторизации матриц, восстанавливают трехмерные координаты только некоторых точек объекта, поэтому возникает задача интерполяции его поверхности между восстановленными точками модели. Восстановление трехмерной сцены по паре изображений, полученных с разных ракурсов, может дать точный трехмерный образ практически для всех точек исходной сцены, но требует высокой точности данных о взаиморасположении камер. Методы получения формы по теням не способны корректно восстанавливать трехмерную информацию об объекте в реальных условиях, когда характер освещения неизвестен и может произвольно изменяться.

Таким образом, наиболее разумным представляется комбинирование различных подходов, предложенное в [6]. Алгоритмы, основанные на факторизации матриц, способны предоставить необходимую информацию о трехмерных координатах камер, их ориентации в пространстве и точности, с которой эти величины известны. В этом случае методы стереосопоставления могут дать плотное восстановление трехмерной поверхности объекта (для каждого пикселя изображения).

Кроме того, интересной для проведения дальнейших исследований является возможность определения и сопоставления соответствующих точек лица в последовательности кадров видео на основе построения модели освещенности сцены.

Литература

1. Вежневец В. Задача восстановления формы объекта по его закраске (shape from shading) [Электронный ресурс] / В. Вежневец // Online журнал «Графика и мультимедиа». – 2003. – Вып. 2. – Режим доступа : http://cgm.graphicon.ru/issue2/Paper_vvp2/
2. William A.P. Smith. Statistical Methods For Facial Shape-from-Shading and Recognition / William A.P. Smith // PhD thesis. – University of York, 2007.
3. Cantzler Helmut. An overview of shape from motion [Электронный ресурс] / Helmut Cantzler // CVonline: The Evolving, Distributed, Non-Proprietary, OnLine Compendium of Computer Vision. – Режим доступа : http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/CANTZLER2/shape_motion.html
4. Манолов А.И. Некооперативная биометрическая идентификация по 3D-моделям лица с использованием видеокамер высокого разрешения [Электронный ресурс] / А.И. Манолов, А.Ю. Соколов // Труды 19-й Международной конференции по компьютерной графике и зрению «ГрафиКон’2009», (Москва, 5 – 9 октября 2009 г.). – М. : МАКС ПРЕСС, 2009. – Режим доступа : http://gc2009.graphicon.ru/en/proceedings
5. Шлянников А.В. Алгоритм генерации трехмерного макета лица по фотографии / А.В. Шлянников // Компьютерная оптика. – 2010. – Т. 34, № 2. – С. 272-277.
6. Свешникова Н.В. Восстановление трехмерных сцен с помощью методов факторизации: принцип работы и оценка погрешностей / Н.В. Свешникова, Д.В. Юрин // Труды 2-й научной конференции «СИМ'2005», (Москва, 19 мая 2005 г.). – Москва, 2005. – С. 207-222.

Literatura

1. Vezhnevec V. Grafika i mul’timedia. Vyp. 2. 2003. http://cgm.graphicon.ru/issue2/Paper_vvp2/
2. William A.P. Smith. Statistical Methods For Facial Shape-from-Shading and Recognition // PhD thesis. University of York. 2007.
3. Cantzler H. CVonline: The Evolving, Distributed, Non-Proprietary, OnLine Compendium of Computer Vision. http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/CANTZLER2/shape_motion.html
4. Manolov A.I. Trudy 19-j Mezhdunarodnoj konferencii po komp’juternoj grafike i zreniju «GrafiKon’2009», (Moskva, 5 – 9 oktjabrja 2009 g.). M.: MAKS PRESS. 2009. Rezhim dostupa : http://gc2009.graphicon.ru/en/proceedings
5. Shljannikov A.V. Komp’juternaja optika. T. 34. № 2. 2010. S. 272-277.
6. Sveshnikova N.V. Trudy 2-j nauchnoj konferencii SIM’2005, (Moskva, 19 maja 2005 g.). Moskva. 2005. S. 207-222.

А.В. Дьяченко

Задача 3D розпізнавання облич: сучасні методи розв’язання

У статті розглянуто сучасний стан проблеми 3D розпізнавання облич, наведено принципи роботи існуючих систем тривимірного розпізнавання. Представлені способи розв’язання ключової задачі у сфері тривимірного розпізнавання – отримання тривимірної інформації про обличчя. Розглянуто поширені методи 3D розпізнавання облич.

A.V. Djachenko

The Review of Modern Methods for Decision of 3D Face Recognition Problem

In the article the state of the art in 3D face recognition is considered, the work principles of three-dimensional existing systems are resulted, the ways for decision of a key problem in area of three-dimensional recognition, i.e. reception of the three-dimensional face information are presented, routine methods for 3D face recognition are considered.

Статья поступила в редакцию 22.06.2011.